

CRFI



IFW16

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/009,874

DATE: 08/04/2004

TIME: 07:19:45

Input Set : N:\efs\10009874\seq_list.txt

Output Set: N:\CRF4\08042004\J009874.raw

```

3 <110> APPLICANT: GOLDBERG, EDWARD B.
5 <120> TITLE OF INVENTION: GENE AND PROTEIN SEQUENCES OF PHAGE T4 GENE 35
7 <130> FILE REFERENCE: 8471-007-228
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/009,874
C--> 10 <141> CURRENT FILING DATE: 2001-12-11
12 <160> NUMBER OF SEQ ID NOS: 8
14 <170> SOFTWARE: PatentIn Ver. 2.0
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1119
18 <212> TYPE: DNA
19 <213> ORGANISM: Bacteriophage T4
21 <400> SEQUENCE: 1
22 atggaaaaat ttatggcoga gtttggacaa ggatatgtcc aaacgccatt tttatcggaa 60
23 agtaattcag taagatataa aataagtata gcgggttctt gcccgctttc tacagcagga 120
24 ccatcatatg ttaaatttca ggataatcct gtaggaagtc aaacatttag cgcaggcctc 180
25 catttaagag tttttgaccc ttccaccgga gcattagttg atagtaagtc atatgccttt 240
26 tcgacttcaa atgatactac atcagctgct tttgttagtt tcatgaattc tttgacgaat 300
27 aatcgaattg ttgctatatt aactagtggg aagggttaatt ttctctctga agtagtatct 360
28 tggttaagaa ccgccggaac gtctgccttt ccatctgatt ctatattgtc aagatttgac 420
29 gtatcatatg ctgcttttta tacttcttct aaaagagcta tcgcattaga gcatgttaaa 480
30 ctgagtaata gaaaaagcac agatgattat caaactattt tagatgttgt atttgacagt 540
31 ttagaagatg taggggctac cgggtttcca agaggaacgt atgaaagtgt tgagcaattc 600
32 atgtcggcag ttggtggaac taatgacgaa attgcgagat tgccaacttc agctgctata 660
33 agtaaattat ctgattataa tttaattcct ggagatgttc tttatcttaa agctcagtta 720
34 tatgctgatg ctgatttact tgctcttgga actacaaata tatctatccg tttttataat 780
35 gcatctaacg gatataattt ttcaacacaa gctgaattta ctgggcaagc tgggtcatgg 840
36 gaattaaagg aagattatgt agttgttcca gaaaacgcag taggatttac gatatacgca 900
37 cagagaactg cacaagctgg ccaaggtggc atgagaaatt taagcttttc tgaagtatca 960
38 agaaatggcg gcatttcgaa acctgctgaa tttggcgtca atggtattcg tgtaattat 1020
39 atctgcgaat ccgcttcacc ccgcatata atggtacttc ctacgcaagc atcgtctaaa 1080
40 actggtaaaag tgtttgggca agaatttaga gaagtttaa 1119
42 <210> SEQ ID NO: 2
43 <211> LENGTH: 372
44 <212> TYPE: PRT
45 <213> ORGANISM: Bacteriophage T4
47 <400> SEQUENCE: 2
48 Met Glu Lys Phe Met Ala Glu Phe Gly Gln Gly Tyr Val Gln Thr Pro
49 1 5 10 15
51 Phe Leu Ser Glu Ser Asn Ser Val Arg Tyr Lys Ile Ser Ile Ala Gly
52 20 25 30
54 Ser Cys Pro Leu Ser Thr Ala Gly Pro Ser Tyr Val Lys Phe Gln Asp
55 35 40 45
57 Asn Pro Val Gly Ser Gln Thr Phe Ser Ala Gly Leu His Leu Arg Val

```

RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/009,874

TIME: 07:19:45

Input Set : N:\efs\10009874\seq_list.txt

Output Set: N:\CRF4\08042004\J009874.raw

```

58      50      55      60
60 Phe Asp Pro Ser Thr Gly Ala Leu Val Asp Ser Lys Ser Tyr Ala Phe
61 65      70      75      80
63 Ser Thr Ser Asn Asp Thr Thr Ser Ala Ala Phe Val Ser Phe Met Asn
64      85      90      95
66 Ser Leu Thr Asn Asn Arg Ile Val Ala Ile Leu Thr Ser Gly Lys Val
67      100      105      110
69 Asn Phe Pro Pro Glu Val Val Ser Trp Leu Arg Thr Ala Gly Thr Ser
70      115      120      125
72 Ala Phe Pro Ser Asp Ser Ile Leu Ser Arg Phe Asp Val Ser Tyr Ala
73      130      135      140
75 Ala Phe Tyr Thr Ser Ser Lys Arg Ala Ile Ala Leu Glu His Val Lys
76 145      150      155      160
78 Leu Ser Asn Arg Lys Ser Thr Asp Asp Tyr Gln Thr Ile Leu Asp Val
79      165      170      175
81 Val Phe Asp Ser Leu Glu Asp Val Gly Ala Thr Gly Phe Pro Arg Gly
82      180      185      190
84 Thr Tyr Glu Ser Val Glu Gln Phe Met Ser Ala Val Gly Gly Thr Asn
85      195      200      205
87 Asp Glu Ile Ala Arg Leu Pro Thr Ser Ala Ala Ile Ser Lys Leu Ser
88      210      215      220
90 Asp Tyr Asn Leu Ile Pro Gly Asp Val Leu Tyr Leu Lys Ala Gln Leu
91 225      230      235      240
93 Tyr Ala Asp Ala Asp Leu Leu Ala Leu Gly Thr Thr Asn Ile Ser Ile
94      245      250      255
96 Arg Phe Tyr Asn Ala Ser Asn Gly Tyr Ile Ser Ser Thr Gln Ala Glu
97      260      265      270
99 Phe Thr Gly Gln Ala Gly Ser Trp Glu Leu Lys Glu Asp Tyr Val Val
100      275      280      285
102 Val Pro Glu Asn Ala Val Gly Phe Thr Ile Tyr Ala Gln Arg Thr Ala
103      290      295      300
105 Gln Ala Gly Gln Gly Gly Met Arg Asn Leu Ser Phe Ser Glu Val Ser
106 305      310      315      320
108 Arg Asn Gly Gly Ile Ser Lys Pro Ala Glu Phe Gly Val Asn Gly Ile
109      325      330      335
111 Arg Val Asn Tyr Ile Cys Glu Ser Ala Ser Pro Pro Asp Ile Met Val
112      340      345      350
114 Leu Pro Thr Gln Ala Ser Ser Lys Thr Gly Lys Val Phe Gly Gln Glu
115      355      360      365
117 Phe Arg Glu Val
118      370
121 <210> SEQ ID NO: 3
122 <211> LENGTH: 8855
123 <212> TYPE: DNA
124 <213> ORGANISM: Bacteriophage T4
126 <400> SEQUENCE: 3
127 taggagcccg ggagaatggc cgagattaaa agagaattca gagcagaaga tggctctggac 60
128 gcaggtggtg ataaaataat caacgtagct ttagctgac gtaccgtagg aactgacggt 120
129 gttaacgttg attacttaat tcaagaaaac acagttcaac agtatgatcc aactcgtgga 180

```

RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/009,874

TIME: 07:19:45

Input Set : N:\efs\10009874\seq_list.txt

Output Set: N:\CRF4\08042004\J009874.raw

```

130 tattttaaag attttgtaat catttatgat aaccgctttt gggctgctat aaatgatatt 240
131 ccaaaaccag caggagcttt taatagcgga cgctggagag cattacgtac cgatgctaac 300
132 tggattacgg ttcatctcgg ttcatatcaa ttaaaatctg gtgaagcaat ttcggttaac 360
133 accgcagctg gaaatgacat cacgtttact ttaccatctt ctccaattga tggtgatact 420
134 atcgtttctc aagatattgg aggaaaacct ggagttaacc aagttttaat tgtagctcca 480
135 gtacaaagta ttgtaaactt tagaggtgaa caggtagctt cagtactaat gactcatcca 540
136 aagtcacagc tagttttaat ttttagtaat cgtctgtggc aaatgtatgt tgctgattat 600
137 agtagagaag ctatagttgt aacaccagcg aatacttata aagcgcaatc caacgatttt 660
138 atcgtacgta gatttacttc tgctgcacca attaatgtca aacttccaag atttgctaatt 720
139 catggcgata ttattaattt cgtcgattta gataaactaa atccgcttta tcatacaatt 780
140 gttactacat acgatgaaac gacttcagta caagaagttg gaactcattc cattgaaggc 840
141 cgtacatcga ttgacggttt ctgatgttt gatgataatg agaaattatg gagactgttt 900
142 gacggggata gttaaagcgcg tttacgtatc ataacgacta attcaaacat tcgtccaaat 960
143 gaagaagtta tggattttgg tgcgaataac ggaacaactc aaacaattga gcttaagctt 1020
144 ccaactaata tttctgttgg tgatactgtt aaaatttcca tgaattacat gagaaaagga 1080
145 caaacagtta aaatcaaagc tgctgatgaa gataaaattg cttcttcagt tcaattgctg 1140
146 caattcccaa aacgctcaga atatccacct gaagctgaat gggttacagt tcaagaatta 1200
147 gtttttaacg atgaaactaa ttatgttcca gttttggagc ttgcttacat agaagattct 1260
148 gatggaaaat attgggttgt acagcaaaac gttccaactg tagaaagagt agattcttta 1320
149 aatgattcta ctagagcaag attaggcgta attgctttag ctacacaagc tcaagctaatt 1380
150 gtcgatttag aaaattctcc acaaaaagaa ttagcaatta ctccagaaac gttagctaatt 1440
151 cgtactgcta cagaaactcg cagaggtaat gcaagaatag caactactgc tcaagtgaat 1500
152 cagaacacca cattctcttt tgctgatgat attatcatca ctccataaaa gctgaatgaa 1560
153 agaactgcta cagaaactcg tagagggtgc gcagaaattg ctacgcagca agaaactaat 1620
154 gcaggaaccg atgatactac aatcatcact cctaaaaagc ttcaagctcg tcaaggttct 1680
155 gaatcattat ctggtattgt aacctttgta tctactgcag gtgctactcc agcttctagc 1740
156 cgtgaattaa atggtacgaa tgtttataat aaaaacactg ataatttagt tgtttcacct 1800
157 aaagctttgg atcagtataa agctactcca acacagcaag gtgcagtaat tttagcagtt 1860
158 gaaagtgaag taattgctgg acaaagtcag caaggatggg caaatgctgt tgtaacgcca 1920
159 gaaacgttac ataaaaagac atcaactgat ggaagaattg gtttaattga aattgctacg 1980
160 caaagtgaag ttaatacagg aactgattat actcgtgcag tcaactcctaa aactttaaat 2040
161 gaccgtagag caactgaaag tttaagtggg atagctgaaa ttgctacaca agttgaattc 2100
162 gacgcaggcg tcgacgatac tcgtatctct acaccattaa aaattaaaac cagatttaat 2160
163 agtactgata gtacttctgt tgttgctcta tctggattag ttgaatcagg aactctctgg 2220
164 gaccattata cacttaatat tcttgaagca aatgagacac aacgtggtac acttcgtgta 2280
165 gctacgcagg tcgaagctgc tgcgggaaca ttagataatg ttttaataac tcctaaaaag 2340
166 ctttttaggta ctaaactctac tgaagcgcaa gaggggtgta ttaaagttgc aactcagctc 2400
167 gaaactgtga ctggaacgtc agcaaatact gctgtatctc caaaaaattt aaaatggatt 2460
168 gcgcagagtg aacctacttg ggcagctact actgcaataa gaggttttgt taaaacttca 2520
169 tctggttcaa ttacattcgt tggtaatgat acagtcggtt ctaccaaga tttagaactg 2580
170 tatgagaaaa atagctatgc ggtatcacca tatgaattaa accgtgtatt agcaaattat 2640
171 ttgccactaa aagcaaaagc tgctgataca aatttattgg atggtctaga ttcatctcag 2700
172 ttcatctgta gggatattgc acagacggtt aatggttcac taaccttaac ccaacaaacg 2760
173 aatctgagtg cccctcttgt atcatctagt actggtgaat ttggtggttc attggccgct 2820
174 aatagaacat ttaccatccg taatacagga gccccgacta gtatcgtttt cgaaaaaggt 2880
175 cctgcacccg gggcaaatcc tgcacagtca atgagtattc gtgtatggg taaccaattt 2940
176 ggcggcggta gtgatacgac ccgttcgaca gtgtttgaag ttggcgatga cacatctcat 3000
177 cacttttatt ctcaacgtaa taaagacggt aatatagcgt ttaacattaa tggtagctga 3060
178 atgccaaata acattaatgc ttccggtttg atgaatgtga atggcactgc aacattcggg 3120

```

RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/009,874

TIME: 07:19:45

Input Set : N:\efs\10009874\seq_list.txt

Output Set: N:\CRF4\08042004\J009874.raw

```

179 cggttcagttta cagccaatgg tgaattcatc agcaagtctg caaatgcttt tagagcaata 3180
180 aacggtgatt acggattctt tattcgtaat gatgcctcta atacctattt tttgctcact 3240
181 gcagccggtg atcagactgg tggttttaat ggattacgcc cattattaat taataatcaa 3300
182 tccggtcaga ttacaattgg tgaaggctta atcattgcc aagggtgttac tataaattca 3360
183 ggcggtttta ctgttaactc gagaattcgt tctcagggtta ctaaaacatc tgatttatat 3420
184 acccggtgcgc caacatctga tactgtagga ttctgggtcaa tcgatattaa tgattcagcc 3480
185 acttataacc agttcccggg ttatttttaa atggttgaaa aaactaatga agtgactggg 3540
186 cttccatact tagaacgtgg cgaagaagtt aaatctcctg gtacactgac tcagtttggt 3600
187 aacacacttg attcgcttta ccaagattgg attacttacc caacgacgcc agaagcgcgt 3660
188 accactcgct ggacacgtac atggcagaaa accaaaaact cttgggtcaag tttgttcag 3720
189 gtatttgacg gaggtaaccc tcctcaacca tctgatatcg gtgctttacc atctgataat 3780
190 gctacaatgg ggaatcttac tattcgtagt ttcttgcgaa ttggtaatgt tcgcattgtt 3840
191 cctgacccag tgaataaaac gggttaaattt gaatgggttg aataagaggt attatggaaa 3900
192 aatttatggc cgagatttgg acaaggatat gtccaaacgc cattttatcg gaaagtaatt 3960
193 cagtaagata taaaataagt atagcgggtt cttgcccgct ttctacagca ggaccatcat 4020
194 atgttaaatt tcaggataat cctgtaggaa gtcaaacatt taggcgcagg ccttcattta 4080
195 agagtttttg acccttccac cggagcatta gttgatagta agtcatatgc tttttcgact 4140
196 tcaaatgata ctacatcagc tgcttttggt agttttcatg aattccttga cgaataatcg 4200
197 aattgttgct atattaacta gtggaaaggt taattttcct cctgaagtag tatcttggtt 4260
198 aagaaccgcc ggaacgtctg cctttccatc tgattctata ttgtcaagat ttgacgtatc 4320
199 atatgtgctt ttttataact cttctaaaag agctatcgca ttagagcatg tttaaactgag 4380
200 taatagaaaa agcacagatg attatcaaac tatttttagat gttgtatttg acagtttaga 4440
201 agatgtagga gctaccgggt ttccaagaaa acgtagtaaa agtggtgagc aattcatgtc 4500
202 ggcagttggt ggaactaata acgaaattgc gagattgcca acttcagctg ctataagtaa 4560
203 attatctgat tataatttaa ttcttgga tgttctttat cttaaagctc agttatatgc 4620
204 tgatgctgat ttacttgctc ttggaactac aaatatatct atccgttttt ataatgcac 4680
205 taacggatat atttcttcaa cacaagctga atttactggg caagctgggt catgggaatt 4740
206 aaaggaagat tatgtagttg ttccagaaaa cgcagtagga ttacgatat acgcacagag 4800
207 aactgcacaa gctggccaag gtggcatgag aaatttaagc ttttctgaag tatcaagaaa 4860
208 tggcggcatt tcgaaacctg ctgaatttgg cgtcaatggt attcgtgtta attatatctg 4920
209 cgaatccgct tcacctccgg atataatggt acttcctacg caagcatcgt ctaaaactgg 4980
210 taaagtgttt gggcaagaat ttagagaagt ttaaattgag ggacccttcg ggttcccttt 5040
211 ttctttataa atactattca aataaagggg catacaatgg ctgattttaa agtaggttca 5100
212 acaactggag gctctgtcat ttggcatcaa ggaaattttc cattgaatcc agccggtgac 5160
213 gatgtactct ataaatcatt taaaatatat tcagaatata acaaaccaca agctgctgat 5220
214 aacgatttctg tttctaaagc taatgggtgt acttatgcat caaaggtaac attttaacgt 5280
215 ggcattcaag tcccatatgc tccaaacatc atgagcccat gcgggattta tgggggtaac 5340
216 ggtgatggtg ctacttttga taaagcaaat atcgatattg ttcatggta tggcgtagga 5400
217 tttaaatcgt catttggttc aacaggccga actggtgtaa ttaatacacg caatggtgat 5460
218 attaacacaa aagggtgtgt gtccgcagct ggtcaagtaa gaagtgtgct ggctgctcct 5520
219 atagcagcga atgaccttac tagaaaggac tatgttgatg gagcaataaa tactgttact 5580
220 gcaaatgcaa actctagggg gctacggctc ggtgacacca tgacaggtaa tttaacagcg 5640
221 ccaaactttt tctcgcagaa tctgcatct caaccctcac acgttccacg atttgaccaa 5700
222 atcgtaatta aggattctgt tcaagatttc ggctattatt aagaggactt atggctactt 5760
223 taaaacaaat acaattttaa agaagcaaaa tcgcaggaa acgtcctgct gcttcagtat 5820
224 tagccgaagg tgaattggct ataaacttaa aagatagaac aatttttact aaagatgatt 5880
225 caggaaatat catcgatcta ggttttgcta aaggcgggca agttgatggc aacgttacta 5940
226 ttaacggact tttgagatta aatggcgatt atgtacaaac aggtggaatg actgtaaacy 6000
227 gaccattggt ttctactgat ggcgtcactg gaaaaatttt cagatctaca cagggttcat 6060

```

RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/009,874

TIME: 07:19:45

Input Set : N:\efs\10009874\seq_list.txt

Output Set: N:\CRF4\08042004\J009874.raw

```

228 tttatgcaag agcaacaaac gatacttcaa atgcccattt atggtttgaa aatgccgatg 6120
229 gcaactgaacg tggcggttata tatgctcgcc ctcaaaactac aactgacggt gaaatacgcc 6180
230 ttaggggttag acaaggaaca ggaagcactg ccaacagtga attctatttc cgctctataa 6240
231 atggaggcgga atttcaggct aaccgtatth tagcatcaga ttcgttagta acaaaacgca 6300
232 ttgcggttga taccgttatt catgatgcc aagcatttgg acaatatgat tctcactctt 6360
233 tggttaatta tgtttatcct ggaaccggtg aaacaaatgg tgtaaaactat cttcgtaaag 6420
234 ttcgcgctaa gtccggtggg acaatttatt atgaaattgt tactgcacaa acaggcctgg 6480
235 ctgatgaagt ttcttgggtg tctgggtgata caccagtatt taaactatac ggtattcgtg 6540
236 acgatggcag aatgattatc cgtaattagc ttgcattagg tacattcact acaaatttcc 6600
237 cgtctagtga ttatggcaac gtcggtgtaa tggcgataa gtatcttgtt ctcgcgaca 6660
238 ctgtaactgg ctgtgcatac aaaaaaactg gtgtatttga tctagttagg ggtggatatt 6720
239 ctgttgcttc tattactcct gacagtttcc gtagtactcg taaaggataa ttgtgctgtt 6780
240 ctgaggacca aggcgcaact tggataatgc ctggtacaaa tgcgtctctc ttgtctgttc 6840
241 aaacacaagc tgataataac aatgctggag acggacaaac ccataatcggg tacaatgctg 6900
242 gcggtaaaat gaaccactat ttccgtggta caggtcagat gaatatcaat acccaacaag 6960
243 gtatggaaat taaccgggtt attttgaaat tggtaactgg ctctaataat gtacaatttt 7020
244 acgctgacgg aactatttct tccattcaac ctattaaatt agataacgag atatttttaa 7080
245 ctaaatctaa taatactgcg ggtctttaat ttggagctcc tagccaagtt gatggcacia 7140
246 ggactatcca atggaacggt ggtactcgcg aaggacagaa taaaaactat gtgattatta 7200
247 aagcatgggg taactcattt aatgccactg gtgatagatc tcgcgaaacg gttttccaag 7260
248 tatcagatag tcaaggatat tatttttatg ctcatcgtaa agctccaacc ggcgacgaaa 7320
249 ctattggacg tattgaagct caatttgctg gggatgttta tgctaaagggt attattgcca 7380
250 acggaaaatt tagagttgtt ggtcaagcg cttagccgg caatgttact atgtctaacg 7440
251 gtttgtttgt ccaagggtgt tcttctatta ctggacaagt taaaattggc ggaacagcaa 7500
252 acgcactgag aatttggaac gctgaatatg gtgctatttt ccgtcgttcg gaaagtaact 7560
253 tttatattat tccaaccaat caaatgaag gagaaagtgg agacattcac agctctttga 7620
254 gacctgtgag aataggatta aacgatggca tggttgggtt aggaagagat tcttttatag 7680
255 tagatcaaaa taatgcttta actacgataa acagtaactc tcgcattaat gccaaactta 7740
256 gaatgcaatt ggggcagtcg gcatacattg atgcagaatg tactgatgct gttcgcccg 7800
257 cgggtgcagg ttcatttgct tcccagaata atgaagacgt ccgtgcgccc ttctatatga 7860
258 atattgatag aactgatgct agtgcatatg ttctattttt gaaacaacgt tatgttcaag 7920
259 gcaatggctg ctattcatta gggactttaa ttaataatgg taatttccga gttcattacc 7980
260 atggcgccgg agataacggt tctacaggtc cacagactgc tgattttgga tgggaattta 8040
261 ttaaaaacgg tgattttatt tcacctcgcg atttaatagc aggcaaagtc agatttgata 8100
262 gaactggtta tatcactggt ggttctggtt attttgctaa cttaaacagt acaattgaat 8160
263 cacttaaaac tgatatcatg tcgagttacc caattggtgc tccgattcct tggccgagtg 8220
264 attcagttcc tgctggattt gctttgatgg aaggtcagac ctttgataag tccgcatac 8280
265 caaagttagc tgttgcatat cctagcgttg ttattccaga tatgcgccc gaaactatca 8340
266 agggtaaaac aagtggctgt gctgttttga gcgctgaggg agatgggtgt aaggctcata 8400
267 gccatagtgc atcggttca agtactgact taggtactaa aaccacatca agctttgact 8460
268 atggtacgaa gggaactaac agtacgggtg gacacactca ctctggtagt ggttctacta 8520
269 gcacaaatgg tgagcacagc cactacatcg aggcattgaa tggtagtggg gtaggtggta 8580
270 ataagatgtc atcatatgcc atatcataca gggcggttg gagtaacact aatgcagcag 8640
271 ggaaccacag tcacactttc tcttttgga ctagcagtgc tggcgaccat tccactctg 8700
272 taggtatttg tgctcatacc cacacggtag caattggatc acatggcat actatcactg 8760
273 taaatagtac aggtaataca gaaaacacgg ttaaaaacat tgcttttaac tatatcgttc 8820
274 gtttagcata aggagagggg cttcggccct tctaa 8855
276 <210> SEQ ID NO: 4
277 <211> LENGTH: 1289

```

VERIFICATION SUMMARY

DATE: 08/04/2004

PATENT APPLICATION: US/10/009,874

TIME: 07:19:46

Input Set : N:\efs\10009874\seq_list.txt

Output Set: N:\CRF4\08042004\J009874.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date